Application/Control Number: 09/998,959
Art Unit: 2655

Docket No.: 2000-0606

REMARKS

Reconsideration and allowance are requested. Claims 1 - 36 are pending and no claims are amended.

Objection to the Specification

The Examiner objected to the specification at pages 14 and 17. Applicant has amended pages 14 and 17 to address the objections set forth by the examiner in the Office Action. Applicant submits the amendments to the specifications address the examiners objections and respectfully requests the Examiner to withdraw the objections.

Objection to the Drawings

The Examiner objected to the drawings as failing to comply with 37 C.F.R. 1.84(p)(5) because they include a reference "108" which the Examiner states is not mentioned in the description. Applicant submits that corrected drawing sheets are not required in reply to avoid abandonment. Applicant has reviewed the drawings and is unable to locate the reference "108" in any of the drawings. Applicant notes that Figures 7a, 7b and 7c include references for 100, 102, 104, 106, 110 and 112. However, there is no reference 108 in the figures. Accordingly, Applicant respectfully submits that the drawings comply with 37 C.F.R. 1.84 and that no corrected drawing sheets are required. However, if necessary, Applicant respectfully requests that the Examiner specifically identify where the reference number 108 is in the drawings and Applicant will readily provide the appropriate correction.

Rejection of claims 1, 2, 5, 25, 30 - 31 and 35 - 36 under 35 U. S. C. 103(a)

The Examiner rejects claims 1, 2, 5, 25, 30 - 31 and 35 - 36 under 35 U. S. C. 103 as being unpatentable over U.S. Patent No. 6, 321, 200 to Casey ("Casey") in view of the Smith et al. article. Applicant traverses this rejection and submits that there is no suggestion or motivation to combine Casey with Smith et al. Casey actually teaches away from combining

its teachings with any other speech-related prior art reference. Furthermore, even if combined, these two references fail to teach each limitation recited in the claims.

We first address claim 1. Applicant will explain how Casey teaches a different subject matter from that recited in claim 1 and how the teachings of Casey cited by the Examiner do not anticipate the claim limitations as set forth in the Office Action. Claim 1 relates to a method of recognizing a received phoneme using a stored plurality of phoneme classes, each of the plurality of phoneme classes comprising class phonemes. The Examiner asserts that this preamble language equates to Casey's disclosure of a method for extracting features from a mixture of signals (the Examiner calls them "a set of phonemes") and cites for support column 4, lines 25-29. However, Casey does not teach extracting features from a set of phonemes. This shall become clear from the discussion below.

The purpose of Casey is to extract features from a mixture of signals. See Abstract. He teaches a system which receives a mixture of signals and filters those signals to produce a plurality of band-pass signals which are in turn windowed to produce a plurality of multi-dimensional observation matrices. Singular value decomposition reduces the dimensionality of the multi-dimensional observation matrices. Column 1 in Casey explains that the invention is applicable to search through a library of video segments that have corresponding audio portions as well. For example, assume a user is looking for a video segment where John Wayne is galloping on a horse while firing his gun. Recognition and identification of such scenes can be achieved through the audio events associated with the video. The audio events in this case would include the rhythmic clop of the galloping horse as well as the percussion resulting from shooting a gun. Therefore, Casey provides a method for extracting features from a mixture of signals which may be acoustic, electric, vibrational or other types of signals. The signals include non-speech audio as mentioned in column 2.

Since the purpose and focus of Casey is to extract features specifically from a mixture of signals, Applicant notes that there is an initial fundamental difference between the

teachings of Casey and the invention recited in claim 1. Claim 1 recites a method of recognizing a received phoneme using a stored plurality of phoneme classes each of the plurality of phonemc classes comprising class phonemes. Nowhere in Casey does hc tcach recognizing the reccived phonemc using a stored plurality of phoneme elasses. Bccause of this fundamental difference, many of the features and limitations of claim 1 discussed below related to class phonemes and phoneme vectors are simply not taught or suggested by Casey.

As part of the training phase, claim 1 recites determining a phonemic vector as a time-frequency representation of the class phoneme. The Examiner asserts that FIG. 1 and column 3, lines 1-10, anticipate this step. Applicant traverses this interpretation because feature 111 in FIG. 1 is simply a reprcsentation of thc band-pass signal for a predetermined frequency range as a result of the processing according to filter 110. Thcrc is no mention in Cascy that this is performed as part of a training phase in a method of recognizing a receive phoneme. Thercfore, Applicant submits that the simple band-pass filters and produccd band-pass signals of Casey cannot be identificd as the same feature as determining a phoneme vector as a time-frequency representation of the class phonemc. Accordingly, this limitation is simply not disclosed by Casey.

Next, the Examiner equates dividing the phoneme vector into phonemes segments with step 120 and columns 3, lines 11 - 13 of Casey. This portion of Casey teaches how each of the band-pass signals is windowed into short 20 millisecond time segments to produce observation matriees. Each matrix includcs hundreds of these segmented samples of the band-pass signals. This subject matter in Casey should not be equated with dividing the phoncmc vector into phonemics scgments. First, Casey makes no reference to phonemic vector. Next, because therc is no reference to a phoneme vector, Cascy makes no refercnce to the dividing of the phoneme vector is into phoneme segmcnts. Thc window taught by Casey simply involvcs dividing each band-pass signal 111 into short time-bascd segments of the band-pass signal. Applicant respectfully submits this subject matter of Casey simply

differs from this step of claim 1 because claim 1 does not involve proeessing a band-pass signal but rather involves dividing a phoneme veetor into phonemes segments.

The Examiner next equates assigning eaeh phoneme's segment into a plurality of phoneme parameters with eaeh windowed band-pass signal of Casey being divided into hundreds of samples. The Examiner changes the term "samples" (used by Casey) in the Offiee Aetion with the term "parameters" whieh used in elaim 1. These two terms, however, are not the same thing. The samples assoeiated with each matrix in Casey are simply hundreds of time-segments associated with the band-pass signal whieh has been windowed. In eontrast, elaim 1 recites dividing eaeh phonemies segment from a phoneme vector into plurality of phoneme parameters. Claim 1 it clearly limited to the eontext of proeessing phonemes whieh is simply not mention by Casey. Accordingly, Applicant respeetfully submits that Casey fails to teaeh this claim limitation.

The Examiner also equates expanding eaeh phoneme segment and plurality and phoneme parameters into an expanded stored phoneme veetor with the expanded veetor parameters with the teaehings and Casey regarding the independent component analysis step whieh produees speetral and temporal features that are expressed as veetors. The independent component analysis relates to estimates of the statistieally most independent component within the segmentation window. The temporal features produced by the independent component analysis are also expressed as vectors and describe the evolution of the spectrum components during the eourse of the segment. See eolumn 3, lines 50 - 53. As taught by Casey, the independent component analysis step is used to reduce the dimensionality of the matrices deseribed above. Aeeordingly, although Casey mentions veetors, Casey simply fails to referenee segments of phonemes or expanding a phoneme segment into a veetor representation with expanded that to veetor parameters as is recited in claim 1. Applieants respeetfully submits that Casey's speetral feature veetors assoeiated with

reducing the dimensionality of the time segment matrices simply differ from the vectors of claim 1.

Claim 1 next recites, as part of the training process for the class phonemes, the step of transforming the expanded stored phoneme vector into an orthogonal form using singular value decomposition. The Examiner equates the observation matrix taught in column 3, lines 10 - 13 as the same as transforming the stored phoneme vector and equates the orthogonal transforming the vector into orthogonal form using singular value decomposition is being taught in column 3, lines 16 - 31. Applicant respectfully traverses this interpretation and asserts that the observation matrices discussed above and discussion of singular value decomposition in Casey do not correlate to the step in claim 1 of transforming the expanded stored phoneme vector in to an orthogonal form.

There are several reasons for this. First, there is no mention of phonemes is being related to the observation matrices or being related to the use of singular value decomposition. Casey has nothing to do with transforming expanded stored phonemic vector. In contrast to claim 1, Casey teaches that singular value decomposition is applied to the observation matrices to produce a reduced dimension matrix. Second, Casey does not teach that the use of matrices in Casey has nothing to do with a training phase for training class phonemes. Accordingly, Applicant submits it is clear that Casey fails to teach the step of transforming a phoneme related vector into orthogonal form.

Applicant notes that the Examiner conceded that Casey fails to disclose the extracted features being used to train class phonemes to recognize class phonemes. This lack of disclosure has been discussed above with regards to the teachings of Casey. However, the Examiner applies the teachings of Smith et al. to fill in the void of Casey's disclosure. Applicant traverses the combination of Casey with Smith et al. and submits that the Examiner has not met his prima facie requirements and that there is no motivation to combine these references.

To establish a *prima facie* case of obviousness, the Examiner must meet three criteria. First, there must be some motivation or suggestion, either in the references themselves, or in the knowledge generally available to one of ordinary skill in the art, to combine the references. Second, there must be a reasonable expectation of success, and finally, the prior art references must teach or suggest all the claim limitations. The Examiner bears the initial burden of providing some suggestion of the desirability of doing what the inventor has done. "To support the conclusion that the claimed invention is directed to obvious subject matter, either the references must expressly or impliedly suggest the claimed invention or the examiner must present a convincing line of reasoning as to why the artisan would have found the claimed invention to have been obvious in light of the teachings of the references." MPEP 2142.

If a proposed modification would render the prior art invention being modified unsatisfactory for its intended purposes, then there is no suggestion or motivation to make the proposed modification. *In re Gordon*, 733 F.2d 900, 221 USPQ 1125 (Fed. Cir. 1984). Further, if the proposed modification of the prior art would change the principle operation of the prior art invention being modified, then the teaching of the reference is not sufficient to render the claims *prima facie* obvious. *In re Ratti*, 270 F.2d 810, 123 USPQ 349 (CCPA 1959). The principles outlined in both these cases are applicable here.

The Applicant submits that one of skill in the art would find no motivation or suggestion to combine Casey with Smith et al. The Examiner asserts that it would obvious to modify Casey's method of extracting features to create template patterns used in speech pattern training in recognition as disclosed by Smith et al. The Examiner concluded that a highly accurate representation of speech could be used in the speech recognizer to increase the chances of correct recognition results.

Applicant notes the there is a difference in the focus and subject matter of each of these references. The MPEP Requires the entire teachings of each prior art reference be

discussed and analyzed to determine whether references in fact teach away from such combination or whether the fundamental principles of operation of the reference would have to be modified to the extent that their combination becomes non-obvious. See MPEP 2143.01. Applicant submits the entire teachings of the prior art do not contain the necessary suggestive power to one of skill in the art to combine these references. Furthermore, as shall be discussed below, Casey introduces his invention specifically to address non-speech audio and thus teaches away from combination with other speech-related references.

Smith et al. teach a template adaptation method in hypersphere word classifier. The purpose of Smith et al. is to improved speech recognition by adapting speech recognition templates in which feedback is provided to assess the accuracy recognition and using that and test utterances. Smith et al. teach modifying the template or moving the template when recognition fails. FIG. 1 of Smith et al. illustrates how a hypersphere representing a template may be moved relative to one another to improve recognition. Smith et al. further teach a template generation process in which the recognizer creates a single composite template for each word in a recognition vocabulary. The template is created in a certain way, which involves, for each word or phrase, generating an average, vector by vector, of all training utterances for that word or phrase. During recognition, the template representing the input or unknown utterance is matched against the stored templates in the recognizer vocabulary. Scores are compared and the best score in the comparison identifies the matching template.

Casey desires to extract features specifically from a mixture of signals using a filterbank to produce a plurality of band-pass signals. Casey mentioned in column 4, line 29 that extracted features from the mixture of signals may be compared against stored data by pattern recognition techniques in order to recognize or identify the components which may include speech phonemes. This is one of a list of uses for the extracted features which further include identifying sound effects, musical instruments, animal sounds or any other corpus based analytical model.

Now, for Casey to be combined with Smith et al., the extracted features taught by Casey which are compared against the stored data would have to relate to a word or phrase which is required by Smith et al. This is because Smith specifically teaches a method based on generating templates for a given word or phrase during recognition and receiving an utterance comprising a word or phrase. An unknown template or input template represents the utterance which is matched against the previously generated recognition templates for the recognizer vocabulary. One of skill in the art may recognize that Casey's teachings regarding extracting features from a mixture of audio signals may be used for determining pattern recognition techniques such as speech phonemes. However, inasmuch as the extracted features are drawn from a plurality of band-pass signals which are generated from a bank of filters, one of skill in the art will recognize that that a word or a phrase may be comprised of a variety of frequencies that when combined create the word or phrase. In Casey's invention, a single word or phrase which spans multiple frequencies would be divided by the filter bank into different bands in the plurality of band pass signals. While Casey discusses the use of classifiers to use pattern recognition techniques for recognizing phonemes, the filter component of Casey may actually make it more difficult to recognize words or phrases or to create an input template comprised of a word or phrase as is taught by Smith et al. This is likely why Casey only mentions speech phonemes (where a word and phrases comprises small snippets of sound, or phonemes, which may be found in a single band-pass signal). Applicant submits that these differences in teachings urges against any finding that one of skill in the art would be motivated to combine their teachings.

Given the complexity of speech recognition processes, Applicant submits that blending a phoneme-based recognition approach as in Casey with a word or phrase template recognition approach taught by Smith et al. would require modification of the fundamental principles of operation of either reference. If Smith et al. were to alter its teachings to focus on phoneme recognition, their entire word/phrase template principles would have to be

altered. Furthermore, if Casey were to incorporate a word/phrase template recognition process, then the filterbank process and feature extraction process would have to be altered to manage words and phrases. Applicant further submits that one of skill in the art would recognize that the fundamental teachings of Smith et al. which are word and phrase template-based would have to be modified to be blended with Casey which focuses on extracting features from a mixture of signals.

The speech phoneme component of Casey is a minor one given the purpose of Casey is to process non-speech signals. When the entire teachings of Casey are studied, Applicant notes that Casey teaches away from combination with speech-related references by establishing that his invention focuses on non-speech sound. The whole teachings of Casey including the title and the abstract focus on the main thrust of his invention, which is extracting features from a mixture of signals in using the filterbank to produce a plurality of band-pass signals. From the beginning of Casey's patent, he clearly avoids speech-based audio in favor of processing non-speech signals from many multi-media sources. As stated in paragraph 1:

> Most prior art acoustic signal representation methods have focused on human speech and music. However, there are no good representation methods for many sound effects heard in films, television, video games, and virtual environments, such as footsteps, traffic, doors slamming, leaser guns, hammering, smashing, thunder claps, leaves rustling, water spilling, etc. These environmental acoustic signals are generally much harder to characterize than speech and music because they often comprise multiple noisy and textured components, as well as higher-order structural components such as iterations and scattering.
> ...
> No such methods exist for "audio" objects, other than when the audio objects are speech.
> ...
> Therefore, there is a need for a robust and reliable representation that can deal with a broad class of signal mixtures. Col. 1, lines 10 - 17, 26-28 and 60 - 63.

Clearly, Casey introduces his invention with a focus on non-speech sound representation, i.e., the "mixture" of sounds that, given previous and ongoing efforts on human speech, focuses on non-speech sound parameterization. Casey essentially teaches that

others are handling the speech aspects of this problem and he is focusing on the non-speech audio objects. In this regard, Casey actually teaches away from exploring other references for speech processing improvements.

The MPEP notes that the mere possibility the references can be combined does not qualify them for being combined under Section 103. Applicant submits that there are a number of reasons why one of skill in the art would not have found motivation for suggestions to combine these references to reject the claims. Therefore, because there is no motivation to combine Smith et al. with Casey, and even if combined, these references fail to teach each limitation of claim 1, Applicant submits that claim 1 is patentable over the cited prior art and in condition for allowance.

Claim 36 is a computer-readable medium claim including similar limitations to those discussed above. For the same reasons set forth above, Applicant submits that claim 36 is patentable over the prior art of record.

Claims 2 and 5 depend from claim 1 and therefore inherit the limitations discussed above. For this reason, Applicant submits that claims 2 and 5 are patentable and in condition for allowance.

Claim 25 recites recognition speech using a database of stored phonemes converted into n-dimensional space. As discussed above, there are many reasons why it is not obvious to combine Casey with Smith et al. Therefore, for this initial reason, Applicant submits that claim 25 is patentable over these references.

Furthermore, Applicant submits that even if combined, these references do not each converting the phoneme into n-dimensional space and comparing the received phoneme to each of the stored phonemes in n-dimensional space. The description of the recognizer section of Smith et al. does not teach converting a phoneme into n-dimensional space. They teach matching "isolated words or phrases against a stored set of templates". Speech recognition is complex, and processing and matching phonemes or snippets of speech differs

from matching templates associated with words and phrases as Smith et al. do. Therefore, Applicant simply notes that because Smith et al. focus on their word/phrase templates exclusively, they do not teach converting a received phoneme into n-dimensional space.

The remaining limitations in claim 25 are not taught by Smith et al. for similar reasons. Since Smith et al. focus on word/phrase templates, they simply fail to teach comparing a received phoneme to each of the stored phonemes in the n-dimensional space or recognizing the received phoneme according to the comparison. Inasmuch as these features are not disclosed by Smith et al. and that there is no motivation to combine the references, Applicant submits that claim 25 is patentable and in condition for allowance.

The Examiner rejects claim 30 as being obvious in view of Casey and Smith et al. Since there is no motivation or suggestion to combine these references, Applicant submits that claim 30 is allowable. Furthermore, claim 30 recites a system for recognizing phonemes using stored phonemes that have been converted into n-dimensional space. The system includes a computer that converts received phonemes into n-dimensional space and wherein the computer compares in the n-dimensional space the received phoneme with each phoneme in the database of phonemes. The Examiner asserts that Smith et al. teach these features. Similar to the discussion above relative to claim 25, Applicant submits that Smith et al. focuses on word/phrase template hyperspheres and simply does not disclose or suggest that their approach applies to phoneme recognition. Therefore, Applicant submits that claim 30 is patentable and in condition for allowance.

Claim 31 depends from claim 30 and recites further limitations therefrom. Accordingly, since this claim is allowable and in condition for allowance.

Rejection of Claims 3-4, 6-7, 16 - 22, 26 - 29 and 32 - 34 Under Section 103

The examiner rejects claims 3 - 4, 6 - 7, 16-22, 26-29 and 32-34 under 35 U. S. C. 103(a) as being a undependable over Casey in view of Smith et al. and further in view of the

Cooper article. Applicant respectfully traverses this rejection submits that these claims are patentable and in condition for allowance.

We first turn to claim 3. Claim 3 recites comparing a distance from the center of the hypersphere of the orthogonal form of the expanded received signal vector with a distance from the center of the hypersphere for each orthogonal form of the expanded stored phoneme vector. Applicant has explained above that there is no motivation to combine Casey with Smith et al. Accordingly, at least for this reason, Applicant submits that claim 3 is allowable.

There are additional reasons why claim 3 is allowable over the cited prior art. Applicant notes that the date of the Cooper reference is 1962 and that accurate speech recognition has been a subject of research for many years. The invention of claim 3 in using distances within the hypersphere for phoneme recognition fills a need that has been studied for a long time. Even given the teachings of Cooper in combination with the other references, the teachings of Cooper have been available for many years and never implemented in phoneme recognition. Accordingly, Applicant submits that the invention solves a long felt, long existing but previously unsolved need to provide improved phoneme recognition. The fact of the Cooper reference has existed for over 40 years without being implemented or suggested in the speech recognition context provides convincing evidence that it is not obvious to utilize such features in phoneme recognition as are recited in claim 3.

The factors that are part of this long-felt need analysis are set forth in MPEP 716.04 and include three components: (1) the need must have been a persistent need that was recognized by those of skill in the art; (2) the long-felt need must not have been satisfied by another before the invention by the applicant; and (3) this invention must satisfy the long-felt need. Applicant submits that these three factors are applicable here. It is well known that accurate speech recognition has been a persistent need in the art for years as can easily be identified publicly by searching IEEE publications and other outlets for research including patent databases. Second, given the further efforts in this important area of research,

Applicant submits that speech recognition has not yet gained the acceptance that it could given the persistent problems in accurate phoneme recognition. Therefore, Applicant submits that this invention is an important step in satisfying the long-felt need for accurate recognition of phonemes in speech recognition.

Furthermore, there are yet other reasons why is cannot be obvious to combine these references. If a fundamental principle of operation of a reference would have to be changed to blend it with another reference, then there is no motivation to combine. This applies to the attempt to blend Cooper with Smith et al. Smith et al. teaches a hypersphere approach where the boundaries are determined by the feature standard deviations and acceptance threshold used by a recognizer. Page 565, column 1. The determination of a match between an input template and one of the stored templates is a "score" calculated according to the equations shown in the "Matching" section. If Cooper's teachings were blended with Smith et al., then Smith et al.'s teachings regarding generating a "score" between the templates to determine a match would have to be altered or abandoned in order to utilize Cooper's hypersphere. His hypersphere requires a comparison of a threshold with the Euclidean distance between the unknown and a fixed point. Page 325. Cooper could not be combined with Smith et al. without an alteration of at least one of the references since they teach different uses of a hypersphere. Since such an alteration would be required to blend these references, there is no motivation or suggestion to combine. Therefore, for the various reasons set forth above, Applicant submits that claim 3 is patentable over the cited prior art references and in condition for allowance.

Claim 4 depends from claim 3 and recites further limitations therefrom. Therefore, Applicant submits that this claim is allowable. Claims 6 and 7 depend from claim 1 and recite limitations therefrom. These two claims are patentable for the same reasons set forth above regarding the lack of motivation to combine Casey with Smith et al. and further for the reason that Cooper is an ancient reference.

Claim 16 recites a method of recognizing speech patterns using stored phonemes. The Examiner combines Casey with Smith et al. and Cooper to reject this claim. Applicant submits that claim 16 is patentable for the reasons set forth above. Namely, there is no motivation or suggestion to combine Smith et al. with Casey and further that there is no motivation or suggestion to combine Cooper with Smith et al. As mentioned above, Casey actually teaches away from any combination with a speech-based reference and the teachings of either Smith et al. or Cooper would have to be fundamentally changed for those references to be blended.

Claims 17 - 22 each depend from claim 16 and recite further limitations therefrom. Accordingly, Applicant submits that these claims are patentable as well.

Claims 26 - 29 each depend from claim 25, discussed above. The Examiner rejects these claims in view of Smith et al., Casey and Cooper. For at least the same reasons set forth above regarding the lack of motivation to combine these three references, Applicant submits that these claims are patentable and in condition for allowance.

Claims 32 - 34 each depend from claim 30 above and the Examiner has rejected these claims in view of Smith et al., Casey and Cooper. Applicant submits that these references cannot be legally combined for the reasons set forth above. Therefore, Applicant submits that these claims are patentable and in condition for allowance.

Rejection of Claims 8 - 15 and 23 -24 Under Section 103

The Examiner rejects claims 8 - 15 and 23 - 24 under Section 103(a) as being unpatentable in view of Casey, in view of Smith et al., in view of Cooper and in view of the Ostendorf article. Applicant traverses this rejection and submits that it has been established above that there cannot be any motivation to combine Casey with Smith et al., or to combine Smith et al. with Cooper. Furthermore, as discussed next, there is no motivation to combine Ostendorf with Casey or with Smith et al.

Ostendorf teaches a stochastic segment model for phoneme-based continuous speech recognition. His approach is introduced specifically as a phoneme-based method where a phoneme is observed as a variable-length sequence of frames where each frame is presented by a parameter vector and where the length of the sequence is random. See Abstract. Already discussed above is the focus of Casey on non-speech audio in extracting features from mixed audio signal sources. Therefore, since Ostendorf clearly focuses on speech recognition, Applicant submits that it would not be obvious to combine Ostendorf with Casey for the same reason that is it not obvious to combine Smith et al. with Casey, namely, that Casey sets forth from the beginning of the disclosure that others are focusing on speech recognition and his goal and purpose is extracting features from non-speech audio.

Furthermore, since Ostendorf teaches a phoneme-based speech recognition method, Applicant submits that it would not be obvious to combine Ostendorf with Smith et al. because Smith et al. focus on a word/phrase-based template approach which as a matter of speech processing differs from the phoneme approach to speech recognition. Clearly, the phoneme recognition algorithm taught on page 1861 of Ostendorf would have to be fundamentally changed or abandoned to blend its teachings with the word/phrase template recognition approach of Smith et al. When Ostendorf does reference words in recognition (page 1867), he builds word models by concatenating phonetic models according to a pronunciation network. This clearly differs from the basic approach of Smith et al. which involves creating a single composite template for each word in a recognition dictionary by using the training utterances for each word. There is no building of a word via concatenation as is taught in Ostendorf.

For the reasons set forth above, one of skill in the art would not find any motivation to combine Ostendorf with Casey because Casey teaches away from extracting features from speech by focusing on non-speech. One of skill in the art would not be motivated to completely alter the word-based approach of Smith et al. to incorporate the phoneme-based

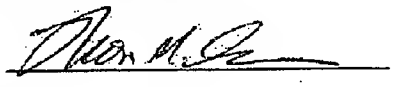
recognition algorithm of Ostendorf. As discussed above as well, Cooper is a very old reference. For all these reasons, Applicant submits that there is simply no motivation or suggestion to combine Ostendorf with the other cited references. Therefore, Applicant submits that claims 8 - 15 and 23 - 24 are patentable and in condition for allowance.

CONCLUSION

Having addressed the rejection of claims 1 - 36, Applicant respectfully submits that the subject application is in condition for allowance and a Notice to that effect is earnestly solicited.

Respectfully submitted,

Date: November 26, 2004

By: 

Correspondence Address:
Samuel H. Dworetsky
AT&T Corp.
Room 2A-207
One AT&T Way
Bedminster, NJ 07921

Thomas M. Isaacson
Attorney for Applicants
Reg. No. 44,166
Phone: 410-414-3056
Fax No.: 410-510-1433